International Journal of Epidemiology 2000;29:793–798

Characteristics of respondents and non-respondents from a case-control study of breast cancer in younger women

M Patricia Madigan,[a] Rebecca Troisi,[a] Nancy Potischman,[b] Donna Brogan,[c] Marilie D Gammon,[d] Kathleen E Malone[e] and Louise A Brinton[a]

Background This study assessed the nature of potential biases by comparing respondents with non-respondents from a case-control study of breast cancer in younger women.

Methods The case-control study was conducted in three regions in the US: Atlanta GA, Seattle/Puget Sound WA, and central New Jersey. An abbreviated interview or mailed questionnaire was completed by willing non-respondents, most of whom had refused participation in the main study.

Results Respondents and non-respondents appeared similar with respect to age, race, relative weight, smoking, family history of breast cancer, number of births, age at first birth, and several dietary items. Compared to non-respondents, case and control respondents were of shorter stature, and reported less frequent consumption of doughnuts/pastries. Respondent cases, compared with non-respondent cases, were more highly educated and more likely to have consumed alcohol regularly; similar but not statistically significant tendencies were observed for controls. Respondent cases experienced menarche earlier than non-respondents. Respondent controls were more likely to have used oral contraceptives than non-respondents; a similar but not statistically significant tendency was observed in cases. Comparisons of crude and simulated relative risks using available non-respondents' data generally showed a low impact of non-response on relative risks in this study.

Conclusions Our results suggest that non-response would not greatly affect relative risk estimates in this study, except possibly regarding height. However, we were limited by the numbers of informative non-respondents and the amount of data collected. Collecting similar information in future studies would be useful, especially since varying methods used to encourage participation may lead to differences in respondents' characteristics.

Keywords Breast neoplasms, epidemiologic study, selection bias, non-response, oral contraceptives, diet, reproductive history, alcohol drinking

Accepted 7 February 2000

[a] Epidemiology & Biostatistics Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.

[b] Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts, Amherst, MA, USA. Formerly at Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.

[c] Rollins School of Public Health, Emory University, USA.

[d] Division of Epidemiology, University of North Carolina, Chapel Hill, NC, USA. Formerly at Division of Epidemiology, Columbia University School of Public Health, USA.

[e] Fred Hutchinson Cancer Research Center, USA.

Reprint requests to: Patricia Madigan, US National Cancer Institute, Division of Cancer Epidemiology & Genetics, Environmental Epidemiology Branch, 6120 Executive Blvd, EPS—MSC 7234, Bethesda, MD 20852-7234, USA.

Case-control studies may be subject to bias when characteristics of respondents and non-respondents are different. Comparisons are needed to assess the nature and direction of biases that researchers are likely to encounter. Several breast cancer studies have compared selected characteristics of respondent controls with non-respondents, or with the general population, and a few characteristics have been compared between case participants and non-participants. However, breast cancer studies comparing a variety of risk factors in *both* case and control respondents and non-respondents are lacking.

We compared characteristics of respondents and non-respondents in a case-control study of younger women which

was designed to include a specific assessment of the relationship of oral contraceptive (OC) use, adolescent diet and alcohol consumption to breast cancer risk. In this study, a short telephone or mailed questionnaire was completed by willing controls and cases who had not participated in the main study.

Methods

The case-control study of breast cancer from which the current study population is drawn is described elsewhere.[1] Briefly, the study was conducted in three geographical areas of the US: the metropolitan areas of Atlanta, GA and Seattle/Puget Sound, WA, and five counties of central New Jersey. All women aged 20–44 years, and who were newly diagnosed with *in situ* or invasive breast cancer 1 May 1990 through 1 December 1992, were eligible for inclusion. In Atlanta the age range extended to 54 years. Potential control subjects were obtained through random-digit dialling.[2] Following informed consent, participants were interviewed in person. The interviewers collected information on: demographic factors, reproductive and menstrual history, contraceptive behaviour, use of exogenous hormones, medical and screening history, anthropometry and physical activity, adolescent diet, alcohol consumption, smoking, occupation, and family history of cancer. A calendar was used to facilitate subjects' recall of OC use. In addition, participants were asked to complete a 100-item recent diet questionnaire, and to consent to a variety of anthropometric measurements. A plan to increase interim response rates was implemented in 1991, including some in-person contacts and the use of selective financial incentives.

Completed interviews were obtained from 2202 (86.5%) of the 2547 eligible cases and 1990 (78.0%) of the 2552 identified eligible control subjects. For cases to be comparable to the control subjects who were identified through telephone sampling, the 29 cases who indicated at interview that they did not have a residential telephone were eliminated leaving 2173 cases available for analysis. Reasons for non-participation included refusals (5.4% physician refusal and 6.5% patient versus 18.6% control subject refusal), death (0.4% versus 0.2%), illness (0.6% versus 0.2%), a move outside the study area (0.6% versus 2.3%), and other miscellaneous reasons (0.2% versus 0.8%). Subjects who did not participate in the main study were asked to complete a short telephone interview or mailed questionnaire. The following groups were not asked: women whose physicians refused participation, women whose physicians informed us that the subject did not wish to be contacted, women whose treatment or membership in a particular medical centre or health maintenance organization prohibited this request, women whose initial refusal was deemed hostile or potentially problematic, and deceased subjects. A total of 51 cases and 168 controls agreed to complete the short questionnaire. The majority of non-respondents who completed the short questionnaire initially did not participate due to subject refusal. The 51 cases include: 29.9% (49/164) of case subject refusals to the main study and 13.3% (2/15) of those whose physicians' consented but who had themselves refused due to illness. The 168 controls include 33.1% (N = 157/475), of subject refusals, 8.6% of those who had moved (N = 5/58), and 40% (N = 6/15) of those who had been unavailable for the main study. The group which completed the short questionnaire comprised 14.4% of non-respondent cases and 29.9% of non-respondent controls.

The non-respondents' questionnaire included questions on OC use, alcohol consumption, recent diet, and recognized breast cancer risk factors. The median phone interview length was 5 min. Most of the questions are comparable to those in the main study, although the question on duration of OC use on the short questionnaire asked about use in terms of years, rather than months.

We compared distributions of characteristics, separately for cases and controls, between respondents and non-respondents, testing for homogeneity using Pearson χ^2 statistics.[3] For height and relative weight (kg/m^2) comparisons, we used reported height for non-respondents and measured height for respondents because they were not asked about height, and for relative weight we used reported weight for all women.

Results

We compared age distributions of respondents and non-respondents, including as 'non-respondents' those who did and those who did not complete the short questionnaires. No differences by age were observed in this group of eligible cases or controls. The remainder of the results are based on comparisons of respondents to the main study with non-respondents who completed the short questionnaire (hereafter referred to as 'non-respondents') since we had no further information on non-respondents who did not complete the short questionnaire.

After stratification by case/control status, respondents and non-respondents appeared similar with respect to race, relative weight, smoking, family history of breast cancer, number of births, age at first birth, and consumption of ground beef, carrots, and eggs (Table 1). Among cases, respondents were more likely to be college graduates than non-respondents (39.8% versus 25.5%, $P = 0.04$); a similar tendency was observed in controls (36.9 versus 30.3%), but was not statistically significant ($P = 0.09$). Both case and control respondents tended to be of shorter stature than non-respondents, and reportedly ate doughnuts/pastries less frequently than non-respondents. The association of shorter height with response showed a similar tendency in both black and white race strata.

Case respondents reported alcohol consumption more frequently, and late menarche less frequently, than case non-respondents; similar tendencies were observed for alcohol in controls, but results did not reach statistical significance. Case respondents reported more frequent consumption of hot dogs than non-respondents, but hot dog consumption was low overall.

Among controls, respondents apparently drank whole milk less frequently than non-respondents, but whole milk consumption was low. Respondent controls were also more likely than non-respondents to have used OC. A similar tendency was observed in cases, but statistical significance was not reached. Because women under 45 had more years of access to OC than older women, and because the parent study of OC focused on women under 45,[1] we examined the association of OC with response among younger women. In women under 45, control respondents were more likely to have used OC than

Table 1 Distribution of characteristics in r espondents and non-r espondents in the W omen's Interview Study of Health (WISH) case-contr ol breast cancer study[a]

	Cases			Controls		
	Subjects (%)			Subjects (%)		
Characteristic	Respondents	Non-respondents	*P*-value	Respondents	Non-respondents	*P*-value
Age (years)			0.56			0.51
<35	12.3	10.8		14.7	13.2	
35–39	22.4	22.7		23.8	22.4	
40–44	41.2	44.2		37.0	36.8	
45–49	12.7	10.2		12.9	15.5	
50–54	11.5	12.2		11.7	12.1	
Education (years)			0.10			0.28
≤High school	27.1	37.3		29.4	35.8	
Post high school[b]	6.8	11.8		8.1	7.9	
Some college	26.3	25.5		25.6	26.1	
College graduate	39.8	25.5		36.9	30.3	
Relative weight (kgs/m^2)			0.88			0.16
≤27.0	74.0	71.4		72.1	79.0	
27.1–≤30	10.1	12.2		11.2	8.6	
≥30.1	15.9	16.3		16.7	12.4	
Ever OC[d] use (age in years)						
20–54	83.4	76.0	0.17	81.4	69.6	0.001
20–44	85.9	76.5	0.12	82.9	73.1	0.007
Ever weekly alcohol intake ≥6 months	59.1	36.7	0.002	54.5	47.2	0.08
Current smoking	14.2	22.5	0.10	21.6	28.0	0.06
Mother or sister with br east cancer	14.3	15.7	0.78	6.3	7.1	0.66
Age at menar che (years)			0.001			0.30
<12	23.6	12.2		23.2	19.9	
12	31.2	22.5		26.3	27.3	
13	27.1	30.6		29.4	29.2	
14	10.1	8.2		11.1	11.2	
≥15	8.1	26.5		10.0	12.4	
No. of births			0.15			0.33
0	22.8	30.6		19.7	13.6	
1	19.3	18.4		18.2	16.7	
2	35.5	28.6		32.4	35.2	
3	15.2	8.2		18.6	21.6	
≥4	7.2	14.3		11.2	13.0	
Nulliparity	22.8	30.6	0.15	19.7	13.6	0.06
Parous	77.2	69.4		80.3	86.4	

Table 1 Continued

	Cases			Controls		
	Subjects (%)			Subjects (%)		
Characteristic	Respondents	Non-respondents	*P*-value	Respondents	Non-respondents	*P*-value
Age at first birth years			0.24			0.51
<1/month	42.3	44.9		45.9	32.1	
1–3/month	41.0	18.4		39.8	40.0	
1–2/week	12.7	24.5		11.1	15.2	
≥3/week	4.0	12.2		3.3	12.7	

a

non-respondents (82.9% versus 73.1%, $P = 0.007$); we observed a similar tendency in cases (85.9% versus 76.5%, $P = 0.12$). In this group of women aged under 45, there were no significant differences by response status, for either cases or controls, in the proportions whose first OC use occurred before age 18, or in the proportions who had used OC for $\geqslant 5$ years, or $\geqslant 10$ years.

Similarly, to assess whether statistically significant associations persisted when subjects' education was considered, we stratified by education (college graduation versus no college graduation) and examined subjects' characteristics. Results were generally consistent with the reported overall tendencies, although statistical power was limited due to the small number of non-respondent cases who completed the short questionnaire. Oral contraceptive use was associated with response among control subjects (81.3% versus 67.8%, $P = 0.001$) who were not college graduates, and marginally with response among cases (83.1% versus 71.1%, $P = 0.05$). A similar tendency was observed among college graduate controls (81.5% versus 72.0%, $P = 0.10$) but not cases (83.9% versus 91.7%, $P = 0.46$).

To assess whether observed differences in OC use, age at menarche, alcohol consumption, and height were likely to have an impact on estimated odds ratios, we computed the relative risks (RR) for these factors using only data from main study respondents, and compared the RR with the RR computed assuming characteristics of non-respondents who completed the short questionnaire accurately reflected the experience of all non-respondents. The crude RR of breast cancer associated with 'ever' use of OC in women under age 45 in the main study was 1.26. If we assume the subset of non-respondents completing the short questionnaire represents OC use in all non-respondents, the crude RR would be 1.37. We did not have detailed information on non-respondents' short term (6–11 months) OC use for comparison with the parent study.[1] The crude RR for ever alcohol use and age at menarche were similar to the simulated RR, although we did not have detailed information on the quantities of alcohol consumed for the non-respondents. The crude RR for a height of 172.7 cm versus a height <157.5 cm was 1.48, while the simulated RR was 2.57.

Discussion

Our results suggest that non-response in this study is not related to age, race, relative weight, smoking, family history of breast cancer, number of births, age at first birth, and several dietary variables. Compared with non-respondents, both case and control respondents tended to be shorter and to report less frequent consumption of doughnuts/pastries. These differences would have little impact on the observed breast cancer RR if the true ratios of diseased to non-diseased subjects by exposure classification were minimally affected.[4] Differences which were statistically significant only among cases, such as age at menarche and alcohol consumption, may have minimal impact on RR because the case response rate was high, and because there were similar tendencies in alcohol consumption in the control group. The difference in OC use among controls would likely be at least partially offset by a similar tendency among cases. Our simulations support this hypothesis. While the simulated RR for the differences in height suggested that a case-control study such as ours could underestimate the effect of height, the height differences could reflect a reporting bias, since height was measured for respondents, but reported for non-respondents.

The differences in dietary factors such as doughnut/pastry and whole milk consumption may reflect different patterns by response status, but could also be chance findings. Although an opposite-direction difference in hot dog consumption was observed among cases, the consumption of hot dogs was low in both respondents and non-respondents, suggesting the observed difference may not be meaningful.

A strength of our study is that we have information on 30% of cases and 33% of controls who refused participation in the main study. However, there are several limitations. We know little about the other two-thirds of the subject refusals, all physician refusals and those who died and most of those who were too ill to participate in the main study. The use of the plan aimed at increasing response rates may limit the generalizability of our comparisons to studies which did not use such a plan, or which used one in a different way, e.g. a plan offering a financial incentive to all subjects. Statistical power to detect differences between respondent and non-respondent cases was limited, and case-control studies including ours generally have no information on people not identified in telephone screening for potential controls. These limitations should be remembered when considering the generally low impact of non-response observed in the simulations.

Previous studies comparing case and control respondents with non-respondents were conducted in various countries, in different age groups, and generally with limited statistical power, and so they provide a limited basis for generalizations. Additionally, methods used to maximize response rates may vary by study and by case/control status, thereby possibly limiting generalizability.

The studies cited in the following paragraphs are case-control studies of breast cancer unless stated otherwise.

Among controls, three breast cancer studies including ours detected little or no age differences between respondent and non-respondent controls: no age differences were detected in Canadian women aged 30–80[5] and a study of US women aged under 55 found control women over 40 slightly more likely to refuse interview.[6] However, a study of Danish women under 70 found respondents younger than non-respondents.[7] Controls in a renal adenocarcinoma study were more likely to be interviewed if they were under 65, and exercised more than an hour per day.[8] A study of laryngeal cancer conducted in Italy found a lower proportion of interviewed subjects among older controls, and among those having less education.[9] The renal cancer study found higher socioeconomic status among control participants compared with non-participants,[8] consistent with the direction of our controls' education, while a Canadian study found no educational differences in comparisons with non-respondent controls.[5]

A study of Norwegian women under 40[10] found OC use lower among 'subject refusal' non-respondents than participating controls, consistent with our finding in control women. Results on parity comparisons are mixed: our study and the Canadian study found that respondents did not differ from non-respondents, but the younger Canadian controls had lower parity than the population;[5] Lund *et al.* reported that Norwegian respondent controls had greater parity than non-respondents;[10]

Morabia *et al.* reported similar parity in Swiss respondents and the population.[11]

Among studies of non-response in cases, Ewertz *et al.* reported that respondents were less likely to be over 59 than non-respondents in the Danish study.[7] As in ours, the Swiss study found similar ages in interviewed and eligible cases,[11] and Wingo *et al.* found little difference by age in comparing interviewed cases age 20–54 who were registered in Surveillance, Epidemiology and End Results (SEER) centres, with cases registered in SEER who were not ascertained and/or not interviewed.[6] Two studies of other cancer sites found similar ages in interviewed and non-interviewed cases: the laryngeal cancer study conducted in Italy,[9] and the renal adenocarcinoma cancer study conducted in Massachusetts.[8]

Consistent with Wingo *et al.*, we found a similar racial distribution among breast cancer case respondents and non-respondents.[6] Not surprisingly, Wingo *et al.* reported that interviewed breast cancer cases were diagnosed at an earlier stage and were more likely to be alive at the most recent follow-up than non-interviewed cases, but that the two groups were similar regarding marital status. Ewertz *et al.* reported that respondent cases and controls were more likely to be married than non-respondents.[7] In the renal adenocarcinoma study, compared to non-interviewed cases, interviewed cases were less likely to be Catholic than Protestant, and less likely to have a history of heavy smoking and heart disease.[8] However, interviewed cases were more likely to have a history of hypertension, and interviewed women were more likely to be in the highest quintile of body mass index, suggesting no clear association of healthy characteristics with case participation.

To conclude, our results provide reassurance regarding the likely impact of response bias on risk estimates in this study, given the response rates achieved and similar tendencies by response status in cases and controls. However, some distortion in RR is possible, especially since our results are based on a subset of non-respondents. Observed differences between respondents and non-respondents suggest that continued consideration of the issue of non-response is needed, especially since varying methods used to encourage participation may lead to differences in respondents' characteristics.

Acknowledgements

This work was supported by National Cancer Institute contracts with Emory University (NO1-CP-95604), New Jersey State Department of Health and Senior Services (NO1-CP-95672), and Fred Hutchinson Cancer Research Center (NO1-CP-95671), and with Westat, Inc. (NO1-CP-95605) to serve as a co-ordinating centre. We wish to thank Shelley Niwa and Eric Mehl for assistance in preparing the data file, and for the many women who agreed to participate in this study.

References

1